Call for Master Thesis

FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

Call for Masters Thesis

Metaphor Identification in Poetry Dataset

Objective of thesis: Metaphor detection and interpretation is a part of natural language processing/natural language understanding task. In metaphor detection, as a first step a sentence is classified into either a metaphor or a literal. If a sentence is classified as metaphorical, the task is then to label the word (unigram) or a sequence of words (n-grams) in the sentence as a metaphor or a literal. The current state of the art algorithms use the context given in the sentence, i.e., if a word is out of context with respect to the rest of the sentence then it indicates this word as a metaphorical expression. For example, *"The alligator's teeth are like white daggers"*, meaning that the teeth of an alligator are white and pointed. Here the word *"daggers"* is used metaphorically and is completely different from the rest of the sentence.

On contrary to the above example, the identification of metaphors in the poetry corpus [1] is much more difficult, because all the words contained in the sentence are comprised of metaphors, e.g.,

“Till it touched the top of heaven.”
The Complete Poetical Works of Henry Wadsworth Longfellow

Following these lines, the objective of this thesis is to annotate a poetry corpus with the state of the art algorithm [2]. for metaphor detection and analyse its output and report the failing examples and the causes of the failures.

[1] https://github.com/aparrish/gutenberg-poetry-corpus
[2] https://github.com/gao-g/metaphor-in-context#data

The thesis will be supervised by **Dr. Mehwish Alam**, details are given below.

Keywords: Deep Learning, Computational Linguistics, Natural Language Processing, Implementation.

Pre-requisites:
- Knowledge of Programming with Python.
- A know how of machine or deep learning and computational linguistics would be a plus.

Please send your CV and questions (if any) to the following e-mail address.

Contact person:
Dr. Mehwish Alam
mehwish.alam@fiz-karlsruhe.de
mehwish.alam@kit.edu

aifb
Institute of
Applied Informatics and
Formal Description Methods
http://www.aifb.kit.edu/

KIT
Karlsruhe Institute of Technology